Information-based functional brain mapping

Nikolaus Kriegeskorte[1], Rainer Goebel[2] and Peter Bandettini[1]

[1]Unit on Functional Imaging Methods, Laboratory of Brain and Cognition,
National Institute of Mental Health, National Institutes of Health,
Building 10, Room 1D80B, 10 Center Dr. MSC 1148, Bethesda, MD 20892-1148, USA

[2]Dept. of Cognitive Neuroscience, Faculty of Psychology, Universiteit Maastricht,
Universiteitssingel 40, 6229 ER Maastricht, The Netherlands

Correspondence should be addressed to:

Nikolaus Kriegeskorte (**niko@nih.gov**)
Unit on Functional Imaging Methods, Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Building 10, Room 1D80B, 10 Center Dr. MSC 1148, Bethesda, MD 20892-1148, **Phone**: 301-594-9195, **Fax**: 301-402-1370

Abstract

We propose a novel method of functional brain mapping based on functional magnetic resonance imaging (fMRI) data. The classical approach is *activation-based* in that it localizes regions that are activated as a whole in one condition as compared to a control condition. Our approach, by contrast, is *information-based* in that it localizes regions whose intrinsic activity pattern contains information about the experimental condition, thus addressing a more general question. The information in local contiguous activity patterns is mapped out by scanning the imaged volume with a spherical "searchlight", whose contents are analyzed multivariately at each location in the brain.

The idea of distributed representations at the spatial scale accessible to fMRI has recently become important in the field of neuroimaging. Our method allows discovery and localization of such representations. We maintain the important idea of mapping the brain for functional regions at the coarse spatial scale, at which interindividual consistency is to be expected. But in contrast to the classical approach, our method is sensitive to information in individually unique fine-grained activity patterns as measured by current high-resolution fMRI (≤2 mm voxel width) but filtered out by conventional analysis in the interest of noise reduction and statistical integration across subjects.

Introduction

Functional brain mapping has evolved from the idea that the brain consists of functionally specialized macroscopic regions. This idea goes back to phrenology, but it received solid scientific support in studies of cytoarchitecture, which established the notion of cortical area, as well as lesion and electrophysiological studies.

In early neuroimaging experiments using positron emission tomography, brain activity was measured at a spatial resolution in the centimeter range. At this resolution the volume elements (voxels) were similar in size to the putative functional regions, so only the spatial-average activity of a region could be studied. In the classical approach to functional brain mapping, therefore, the experiment is designed to activate a functional region as a whole. The region is then localized by computing an activation statistic for each location of the imaging volume and thresholding the resulting statistical map. We refer to this approach as *activation-based*.

With the advent of fMRI, spatial resolution increased. Standard functional measurements were performed using voxel widths of about 4 mm in each dimension. Although a typical functional region at this resolution is covered by multiple voxels, standard fMRI analysis to this day has remained true to the activation-based approach, in which a region is assumed to become active as a whole. This manifests itself in the widespread investigation of the spatially averaged activity for regions of interest. Event-related average time courses and bar graphs depicting the activity across conditions, for example, reflect a region's spatially averaged activity.

More importantly, the assumption that functional regions extended across multiple voxels will become activated as a whole plays a key role in statistical inference at the level of whole maps in several established methods, including the widespread statistical parametric mapping (1, 2, 3, see also: 4). The extended-activations assumption motivates the spatial smoothing of the data, which is standardly performed. Spatial smoothing accentuates extended activations by removing the "salt-and-pepper" fine-structure of the activity patterns,

which is treated as noise. As a positive side effect, the resulting reduction of the data's spatial complexity alleviates the multiple-comparisons problem: With smoothing, controlling the false-positives rate when performing a statistical test at each location of a functional volume will require weaker correction. This entails greater statistical power for detection of extended activations. Furthermore, spatial smoothing helps integrate analyses of different subjects coregistered in Talairach space (5), where corresponding functional regions can be off by many millimeters between subjects. In a typical group analysis, data are spatially smoothed by convolution with a Gaussian kernel of 8-mm full width at half maximum (FWHM).

Although smoothing greatly reduces the information content of the data, the local combination of signals it provides is necessary. If smoothing is omitted in a standard voxelwise univariate fMRI analysis, statistical sensitivity suffers and fewer voxels are detected (Fig. 2). Upon lowering the threshold, the activation maps show salt-and-pepper patterns, which are hard to distinguish from noise, inconsistent across subjects and impossible to report verbally. Nevertheless these fine-scale patterns of weak effects may contain neuroscientifically relevant information.

The amount of information removed by smoothing fMRI data to the scale of functional regions increases with growing spatial resolution of the measurement. Today, a voxel width of 2 mm in each dimension is robustly achievable with standard clinical scanners at 3-Tesla field strength. Using high magnetic field strengths and advanced methods such as parallel imaging (6), fMRI spatial resolution is invading the submillimeter range. Although high-resolution fMRI maps are somewhat compromised by hemodynamic blurring, distortion and vascular artefacts, the mapping of human ocular dominance columns has demonstrated that submillimeter neuronal activity patterns are reflected in the fMRI signal (7, 8, 9, 10, see also 11). We are able to measure activity at a much finer scale than before. This mesoscopic view of brain activity will form the centerpiece in bridging the gap between neuroimaging and invasive electrophysiological recording. However, the neuroscientific exploitation of this information poses new conceptual, experimental and statistical challenges.

The activation-based approach discards the high-resolution information fMRI now provides. Should fine-scale fMRI activity patterns be thought of in terms of distributed representation? How can informative patterns be sensitively detected, distinguishing them from noise of similar salt-and-pepper appearance? How can data be related between subjects when the available common spaces (e.g. Talairach space, 5) lack precision and, more fundamentally, the fine-scale activity patterns, like fingerprints, may be unique to each individual? We propose to approach these challenges by abstracting from the actual patterns of activity in a local neighborhood and considering the *information* they convey about the experimental condition.

Information contained in distributed fMRI activity patterns has been analyzed for extended predefined regions (12), heuristically chosen discontiguous sets of voxels (13, 14) and global patterns (e.g. 15, 16). Here we show how such information can be used to continuously map a functional volume. Instead of searching the functional volume for regions whose spatially averaged activity changes across conditions (activation-based approach), we search for regions whose intrinsic activity pattern changes. Our approach is *information-based* in that it localizes regions containing information about the experimental condition. Information-based mapping is a generalization of activation-based mapping in that it is sensitive to both regions activated as a whole and regions whose spatial-mean activity stays constant, but whose fine-grained activity pattern changes with the experimental condition.

An implementation of information-based functional mapping

The statistical combination of signals from all voxels within a region can afford a reduction of noise and greater statistical power. When the target of the analysis is to find activated regions, local signals can be combined by local averaging, i.e. by smoothing of the data. Here the target of the analysis is to find informative regions and smoothing would remove information in the spatial fine-structure of the data. We therefore combine local signals using a multivariate statistic that compares the activity *patterns* between the conditions.

Spherical multivariate searchlight

In order to obtain a continuous map, in which informative regions are marked, we move a spherical multivariate "searchlight" through the measured volume. The searchlight is centered on each voxel in turn. A searchlight of 4-mm radius, for example, will comprise 33 voxels of 2-mm width in each dimension (see Fig. 1D). To combine the signals from all the voxels falling into the searchlight, a multivariate effect statistic is computed at each location. For each voxel in the volume, the resulting map shows how well the multivariate signal in the local spherical neighborhood differentiates the experimental conditions.

Multivariate effect statistics

In the present implementation, we focus on information carried by the *spatial* pattern of activity within each region. Furthermore we assume that the activity pattern, though covered in noise, is replicable and can, thus, be estimated as the average across trials within each condition (or a least-squares fit, for each voxel, of one predictor time course per condition). These assumptions restrict the types of activity-pattern information that can be detected, excluding information carried by the temporal activity pattern and the shape of the statistical distribution of patterns associated with each condition. However, in a first step towards generalization of the much more restricted activation-based approach, it seems wise to retain some assumptions to improve the stability of the estimates.

We consider the simple case of two conditions. What statistic should be used to represent the difference between the two associated activity patterns within the searchlight?

Average absolute t value. One approach would be to first perform a conventional linear-regression contrast analysis to obtain a t value for each voxel. The t value represents the activity difference between the two conditions in that voxel. One may then average the absolute t values within the searchlight to obtain a measure of the difference between patterns. Taking the *absolute* values is crucial here, since opposite effects would cancel out if the raw t values were averaged. For instance, the average t value might be zero if activity in half the voxels is greater during condition A and activity in the other half greater during

condition B – a case in which the activity pattern contains a lot of information about the condition. The average absolute t value within the searchlight combines local signals without obliterating the spatial fine-structure. However, it does not take the covariance structure of the noise into account.

Mahalanobis distance. A natural choice that accounts for noise covariance is the Mahalanobis distance, a multivariate generalization of the t value. To obtain a t value, the estimated difference between the two activity levels is divided by its estimated variability. To obtain the Mahalanobis distance, the estimated difference vector between the two activity patterns is normalized by the estimated multivariate noise covariance (Eq. 4).

Mancova-p value. In order to determine whether the local activity pattern within the searchlight differs significantly between the two conditions, we perform a standard multivariate analysis of covariance (mancova), which provides a p value for the multivariate contrast between the conditions. The test involves removing the dimension defined by the contrast from the model space, fitting the full and the reduced model and determining the extra-sums-of-squares-and-products matrix associated with the contrast. This matrix is related to the error sums-of-squares-and-products matrix and inference is performed via Wilk's Λ, Bartlett's statistic and the χ^2 distribution (17).

Mathematical details

Estimation of spatial activity patterns. To estimate the spatial activity pattern elicited during each condition, we fit a linear model to each voxel's time course by ordinary least squares. The model comprises a hemodynamic response predictor (Boynton et al., 1996) for each condition and, optionally, additional predictors to model artefactual components such as trends, head-motion effects and baseline shifts between measurement runs. For the volume as a whole, the linear model can be stated as follows:

$$Y = XB + E, \quad (1)$$

where Y is the time-by-voxel data matrix, X is the time-by-predictor design matrix, B is the predictor-by-voxel matrix of beta weights, and E is the time-by-voxel matrix of errors. Least-squares minimization of $sum(diag(E^T E))$ leads to

$$\hat{B} = (X^T X)^{-1} X^T Y, \quad (2)$$

where $\hat{B}$ is the predictor-by-voxel matrix of beta estimates. Each row of $\hat{B}$ that represents an experimental condition (as opposed to trend or head motion) contains an estimated spatial activity pattern.

Multinormal noise model. For the contents of the searchlight at each position, we assume the errors (one spatial pattern of errors for each time point) to have a multinormal distribution in the space spanned by the searchlight voxels. At each searchlight position, the error covariance matrix Σ is estimated as $\hat{\Sigma} = (\hat{E}^T \hat{E})/(n - np)$, where $\hat{E} = Y - X\hat{B}$ (with the voxel set restricted to the searchlight voxels), n is the number of time points and np is the number of predictors in the design matrix including nuisance and confound-mean predictors. For a given condition represented by predictor p, the spatial-activity-pattern estimates have a multinormal distribution characterized by the covariance matrix

$$\Sigma_p = \Sigma \cdot (X^T X)^{-1}_{p,p}. \quad (3)$$

Mahalanobis distance. Consider two conditions and their associated spatial-activity-pattern estimates a_1 and a_2 within the searchlight. The rows of $\hat{B}$ corresponding to the two conditions provide the estimates a_1 and a_2. The Mahalanobis distance is defined as

$$\Delta^2 = (a_2 - a_1)\tilde{\Sigma}^{-1}(a_2 - a_1)^T, \quad (4)$$

where $\tilde{\Sigma}$ is the covariance matrix estimate describing the variability of the activity-pattern estimates a_1 and a_2. For $\tilde{\Sigma}$ to describe the variability of both a_1 and a_2, the design should be

symmetrical with respect to the two conditions. As Eq. 3 shows, $\tilde{\Sigma}$ is a scaled version of the noise dispersion matrix $\hat{\Sigma}$. For localization of informative regions, the scaling of $\tilde{\Sigma}$ is irrelevant and simply using $\hat{\Sigma}$ gives a diagnostically equivalent map of Mahalanobis distances.

Validation of the proposed method

Simulated fMRI data

The classical activation-based approach is based on the assumption that effects are spatially smooth (i.e. in the low spatial-frequency band). In this paper we explore how fMRI activity-pattern information *in any spatial-frequency band* can be utilized for the localization of functional regions. We therefore simulate effects with no spatial-frequency bias: The effects have equal power in all spatial-frequency bands, containing information in their locally averaged component as well as in their spatial fine-structure.

We simulate a time course of three-dimensional volumes as might be obtained in a slow event-related fMRI experiment with two conditions. Each condition is associated with a spatial pattern of response, which fades in and out according to a linear model of the temporal structure of the hemodynamic response. The two spatial patterns are composed of values drawn independently from a Gaussian distribution and confined to realistically shaped regions of four different sizes (Fig. 1A and 1B), which are embedded in a functional volume. The signal data is added to spatiotemporal noise with a realistic degree of correlation between adjacent voxels.

Results of simulation

In order to assess how well different local descriptive statistics distinguish between voxels belonging to informative regions and surrounding noise voxels, we use receiver-operating characteristics (ROC, Fig. 1). Since the effect regions are known for the simulations, the proportion of correctly detected voxels among all region voxels (the sensitivity) and the proportion of correctly rejected voxels among all non-region voxels (the specificity) can be

determined for all possible thresholds. An ROC shows the resulting relation between sensitivity and specificity (Fig. 1A). As a measure of the diagnosticity of the different local statistics, we use the area under the ROC (Fig. 1B and 1C), which reflects to what extent high sensitivity and specificity can simultaneously be achieved. More precisely, the area under the ROC represents three intuitively meaningful quantities: 1) the average sensitivity across all specificities, 2) the average specificity across all sensitivities and 3) the probability that upon drawing a random region voxel and a random non-region voxel, the region voxel will have the higher effect statistic. If the area under the ROC is 1, the statistic allows perfect classification of region and non-region voxels by choice of an appropriate threshold. If the area under the ROC is 0.5, then the statistic does not order region and non-region voxels better than chance. For the parameter space we tested in our simulation, we observe the following.

(1) Information-based mapping using the Mahalanobis distance dominates the other techniques. With an appropriate searchlight radius, Mahalanobis mapping (red curves in Fig. 1A-C) is most sensitive at all specificities for all region sizes, functional-contrast-to-noise ratios and degrees of data smoothing. Best detection performance is achieved using the intermediate-size searchlight (4-mm radius), except when the regions are large (90, 270 voxels) and the functional-contrast-to-noise ratio is low. In that case a slightly larger searchlight (5-mm radius) yields slightly better diagnosticity.

(2) The activation-based mapping performs worse than the information-based approaches. Univariate t mapping (black curves in Fig. 1A-C) performs worse than all information-based statistics (colored curves) for all combinations of region size (small or large) and functional contrast (high or low) for unsmoothed as well as moderately smoothed data. This is because the univariate approach fails to benefit from local spatial combination of signals.

(3) Spatial smoothing of the data monotonically degrades diagnosticity of all statistics. For all combinations of region size and functional-contrast-to-noise ratio, smoothing degrades performance of all statistics tested including the univariate t value. This reflects the fact, that effects in our simulation were equally strong in all spatial-frequency bands up to the Nyquist limit imposed by voxel size. Data smoothing filters out the high-spatial-frequency component of the effect patterns and, thus, reduces the overall effect energy.

(4) The benefit of spatial signal combination within the searchlight is greater for larger regions. For univariate mapping, detection is equally poor independent of region size. The information-based techniques using a spherical searchlight of 2-mm radius or larger (colored curves) perform markedly better than the activation-based univariate mapping (black curves) for small regions and the benefits of spatial combination of signals within the searchlight are even greater for larger regions. This holds for low and high functional contrast and for all searchlight sizes (except the degenerate 1-voxel searchlight, for which both information-based approaches become diagnostically equivalent to the activation-based univariate t mapping).

(5) A searchlight of 4-mm radius yields optimal or near-optimal detection performance. When the size of the spherical searchlight is very inappropriate to the size of the region (small region scanned with large searchlight or vice versa), detection performance suffers. The optimal searchlight radius depends on the shape and size of the regions to be detected. But for all region shapes and sizes simulated, the 4-mm searchlight performed optimally or near-optimally. Detection, thus, is not strongly dependent on an exact match between region and searchlight.

(6) Even a small searchlight affords a large increase of detection performance. Spatial signal combination with even a 2-mm-radius searchlight comprising only seven voxels drastically

improves performance in comparison to univariate t mapping (except for highly smoothed data with performance near chance level).

(7) A noise model can improve detection performance. The Mahalanobis distance outperforms the average absolute t value for any given searchlight size. This is because the Mahalanobis distance is based on a multinormal noise model, which picks up on the realistic correlation present in the simulated noise. The noise model helps separate signal from noise.

Simulation details

A simulated data set was generated by Eq. 1, where X embodies the temporal, B the spatial structure of the signal and E the spatiotemporal noise.

Temporal structure of the signal. We simulate a two-condition slow event-related experiment. Events last 500 ms and their onsets are separated by 16 s. There are 20 events per condition. The entire simulated experiment, thus, lasts 10 min and 40 s. The condition sequence is random. The temporal resolution is one volume per 2 s. The time course of the fMRI signal associated with each condition is simulated by a linear hemodynamic-response model (18) assuming instantaneous rectangular neural responses to each simulated event.

Spatial structure of the signal. The simulated functional volume consists of 9 slices of a size of 128 by 128 voxels. The voxels are assumed to be 2 mm wide in each dimension. Along one dimension of the slice, we vary the size of the regions in four levels (10, 30, 90, 270 voxels). Along the other dimension, we vary the functional-contrast-to-noise ratio in four levels (0.1, 0.2, 0.3, 0.4). Each slice is accordingly subdivided into 4 by 4 subblocks of dimensions 32 by 32 by 9 voxels. Within each subblock representing the smallest and second smallest region size, the region shapes are repeated four times, to provide sufficient data for the case of small regions. The functional-contrast-to-noise ratio is defined as the

spatial average within the effect region of the absolute activity level at the maximum of the hemodynamic response divided by the temporal standard deviation of the noise. The shapes of the regions (exactly as shown in Fig. 1) were created using a region growing process prioritized by a Gaussian random field. In order to obtain realistically compact randomly shaped regions whose thickness does not grossly exceed what one might expect of functional regions in the human cortex, the region growing was spatially biased by means of a disc-shaped pedestal embedded in the Gaussian random field.

Spatiotemporal structure of the noise. To match the correlation found between the residual time courses of neighboring voxels in real fMRI data, the noise is generated by slight spatial smoothing of spatiotemporal Gaussian white noise with a Gaussian kernel of 2.35-mm FWHM. This leads to an adjacent-voxel linear correlation across time of 0.25, approximately matching that found in the residuals of our real fMRI data (see below), which were acquired at the same resolution as simulated here.

Real fMRI data

The simulated focally distributed effects were better detected by information-based mapping. But do such effects actually occur in fMRI data? In order to test this, we analyzed real fMRI data using activation- and information-based techniques. If effects are predominantly spatially smooth, then the activation-based approach should find more voxels, because its implicit smoothness assumption should make it more sensitive to such effects. If effects with a substantial fine-grained component are widespread (as in the simulation), information-based mapping should find more voxels, because it is sensitive to effects in all spatial frequency bands.

We performed blood-oxygen-level-dependent (BOLD) fMRI measurements with voxels of a size of 2×2×2 mm^3 during perception of face and house images. For the information-based analysis, we chose a searchlight radius of 4-mm, because this radius yielded the best

detection performance in the simulation. For the activation-based analysis, the data were spatially smoothed with a spherical kernel of 4-mm radius, in order to exactly match the range of spatial combination of signals between the two approaches. (The resulting degree of smoothing is comparable to that obtained with the conventional choice of a Gaussian of 8-mm FWHM.)

All maps were thresholded such that the average false discovery rate (i.e. the proportion of falsely marked voxels among all marked voxels) was 5% or smaller (19, 20).

Results of fMRI analysis

We focus on the contrast between activity evoked by images of faces and houses. The activation-based map (Fig. 2B) highlights regions more strongly active during face than house perception and vice versa. The information-based map (Fig. 2C) highlights regions whose activity pattern distinguishes the two categories. Qualitatively, we observe that the information-based mapping marks more voxels, highlighting extended swaths of cortex, including most of the category-selective regions (21, 22) found by the activation-based mapping. A quantitative comparison of the maps (Fig. 2D) shows three sets of voxels: The set marked only by information-based mapping (red), the set marked by both information- and activation-based mapping (yellow) and the set marked only by activation-based mapping (green). We consider the three sets in the order of their size.

(1) The information-based mapping marks a large number of voxels not marked by the activation-based mapping. Of all voxels marked in either map, 45% are marked only in the information-based map (red in Fig. 1D-F) on average for the group of 11 subjects analyzed. This indicates that there are many regions containing distributed category information, a result that is consistent with the findings of Haxby et al. (12, see also 23). The activation-based mapping does not detect these regions, because the category information is lost when the data are smoothed. Note, however, that performing conventional univariate analysis without smoothing (Fig. 2A) does not lead to the detection of these regions, either. This is

because the effects are weak in single voxels and therefore only detected with local spatial combination of signals.

(2) A substantial number of voxels is marked by both mapping techniques. Of all voxels marked in either map, 38% are marked in both maps (yellow in Fig. 1D-F) on average for the group. The overlap in marked voxels reflects the sensitivity to extended activated regions shared by the two approaches.

(3) The activation-based mapping marks a small number of voxels not marked by the information-based mapping. Of all voxels marked in either map, 17% are marked only in the activation-based map (green in Fig. 1D-F) on average for the group. The effects detected in these voxels are likely to be extended homogeneous but weak activations, to which the activation-based approach is more sensitive because of its implicit assumption of extended activations. That the information-based approach fails to detect these voxels represents the cost of broadening the focus of sensitivity to a more general class of effects.

Details on fMRI experiment and analysis

Stimuli, design and task. Subjects continually fixated while viewing images of faces and houses. We used a rapid event-related design with a basic trial duration of 3 s. The stimulus sequence was optimized by a method based on a genetic algorithm (24). Each image was presented for 400 ms. The basic set of stimuli consisted of four grayscale photographs, depicting two faces and two houses. In order to enforce and monitor attentive viewing, we asked subjects to perform an anomaly-detection task. They detected subtly altered versions of the four original images. Anomalous trials constituted 12% of all trials and were excluded from the analysis. Subjects pressed buttons placed underneath their right and left index fingers on regular and anomalous trials, respectively.

Subjects. Eleven subjects between 18 and 30 years of age participated. All had normal or corrected-to-normal vision. Five of them were female, six male. All received information about magnetic resonance imaging and gave their informed consent. The experimental techniques used in this study and the consent form were approved by the ethical committee CWOM of the Academisch Ziekenhuis (university hospital) associated with the Katholieke Universiteit Nijmegen (The Netherlands).

Measurements. We measured 15 transversal functional slices with a Siemens Magnetom Trio scanner (3 Tesla) using a single-shot gradient-echo echo-planar-imaging sequence. The pulse-sequence parameters were as follows: in-plane resolution: 2×2 mm^2, slice thickness: 2 mm (no gap between slices), time between functional volumes: 1500 ms, slice acquisition order: interleaved, field of view: 256×256 mm^2, acquisition matrix: 128×128, time to echo: 32 ms, flip angle: 75 deg. The analysis for each subject is based on one measurement run lasting 14.8 min. Imaging was performed at the Donders Centre for Cognitive Neuroimaging (Nijmegen, The Netherlands).

Preprocessing. The fMRI data sets were subjected to a series of preprocessing operations: 1) slice-scan-time correction, 2) head-motion correction, 3) removal of temporal drifts of frequencies below 0.009 Hz, 4) selection of brain voxels for the Benjamini-and-Hochberg procedure (19, 20). Steps 1-3 were performed using the BrainVoyager 2000 software package (version 4.8). Step 4 was performed using custom software developed in Matlab. The brain mask (matching the anatomical background shown in Fig. 2) was computed by thresholding a smoothed version of the pseudoanatomical volume provided by the temporal average of the functional volumes.

Conventional aspects of the analyses. In both, the activation- and the information-based analysis, we assumed an instantaneous rectangular neural response to the stimuli and obtained hemodynamic response predictors using a linear model (18). The design matrix

included these predictors along with six head-motion-parameter time courses. The activation-based analysis consisted in standard univariate linear regression and contrast analysis. All activation- and information-based analyses were performed using custom software developed in Matlab.

Discussion

We propose information-based functional brain mapping, an alternative to the classical activation-based approach. The core idea is to localize functional regions carrying a particular type of information by scanning the entire volume with a multivariate searchlight. By simulation, we show that information-based mapping is more sensitive to focally distributed effects with equal power in all spatial-frequency bands than activation-based mapping. By experiment, we show that focally distributed effects better detected by information-based mapping actually occur in fMRI data.

What experiments are amenable to information-based mapping?

The methods described here can be applied in the context of any fMRI experiment. Block- as well as slow and rapid event-related designs can be analyzed by the methods defined. However, the added value of information-based analysis depends on the neuroscientific questions to be addressed and the experimental design chosen. If the experiment is designed to activate a particular functional region as a whole, an activation-based mapping may be conceptually more natural as well as a more sensitive means of localizing the region. By contrast, experiments that target focally distributed activity patterns, often interpreted as serving a *representational function* in the brain, are naturally suited for information-based mapping. The technique is not restricted to vision or perception. It can be applied in all domains including motor control and higher cognitive function.

Information theory and the generality of multivariate effect measures

Since the method targets activity-pattern information, it is natural to use information-theoretic measures to quantify effects. For example, one can use an estimate of the mutual

information between the experimental conditions and the concomitant local activity patterns. We do not present mutual information estimates in our analysis here, because we are primarily concerned with the localization of informative regions.

Inspired by the generality of the concept of information, the method may be extended in the future to detect differences between spatio*temporal* activity patterns and between other aspects of the pattern distributions than their centroids. However, broadening the focus of an analysis to a more general class of effects comes at a cost in sensitivity. We feel that the modest generalization of the classical activation-based approach proposed here is realistic for fMRI data and well worth its cost.

Spherical or cortex-patch searchlight

Information-based mapping statistically combines local signals within a multivariate searchlight. It is, thus, most sensitive to information in contiguous macroscopic regions well sampled by the searchlight. For optimal joint sampling of informative voxels, the radius of the spherical searchlight should reflect the size and compactness of the regions. However, our simulation shows that a radius of 4 mm yields near-optimal performance for small as well as large regions of realistic shape. In localizing cortical information, using an explicit representation of each subject's cortical sheet (25, 26, 27) and replacing the spherical searchlight by a cortex patch promises to improve the joint sampling of informative signals and enhance sensitivity.

Group analysis

A group-analytical extension of the information-based approach to functional mapping has the potential to address an important challenge functional brain mapping is currently faced with: While human fMRI already operates at a resolution of 1-2 mm, the spatial reference frames relating locations in different individual brains have a much lower spatial precision. Information-based group analysis should therefore follow a two-scale approach. At the fine spatial scale of millimeters, activity patterns are assumed to be unique to each individual and therefore analyzed separately for each subject as described here. At the coarse spatial scale

of centimeters, single-subject multivariate effect statistics are combined to increase statistical power and obtain a group-statistical summary of the individual results.

Acknowledgements

The authors thank Ziad Saad and Leslie Ungerleider for helpful comments on the manuscript.

Figure legends

Fig. 1: **Simulated fMRI data**

A-C show the performance of activation- (black lines) and information-based mapping (colored lines) at detecting focally distributed effects (for color coding, see **D**). **A** shows ROCs for the case of unsmoothed data. **B** shows the effect of spatial smoothing of the data. The vertical axis represents the area under the ROC. For the included case of no smoothing (i.e. FWHM=0), the areas under the ROC (marked as circles) correspond to the ROCs shown in A. Note that smoothing degrades performance for all techniques. For the crucial case of unsmoothed data, **C** summarizes the essential results by visually relating the detection performances afforded by the different techniques for small and large regions and low and high functional-contrast-to-noise ratio. The searchlights yielding optimal performance in each case are shown in gray (4- or 5-mm radius). The circles in C replicate the circles in B reflecting the areas under the ROCs shown in A. In **A** and **B**, the line thickness measured vertically is 2 standard errors of the mean obtained by repeating the simulations and analyses 40 times with fresh noise. The shapes of the regions shown in green in A-C are exactly those used in the simulations. **D** illustrates the color coding in A-C and shows the searchlights used.

Fig. 2: **Real fMRI data**

Comparison of activation- and information-based mapping techniques in application to real fMRI data. **A-C** show different analyses of a single subject's data. **A and B** show univariate two-sided t maps contrasting the strength of activity during perception of face and house images. The color scale linearly reflects the t value for the contrast between faces and houses (see color bar) for voxels above the FDR threshold. Yellow and cyan voxels have absolute t values close to the maximum absolute t value found in the map. **A** shows the t map for unsmoothed data, **B** for smoothed data. Note that smoothing increases the number of voxels marked. **C** shows the information-based map of p values from mancova performed at each location for the contents of a spherical searchlight. The color scale linearly reflects the p value (see color bar). **D** shows a direct comparison of the voxel sets marked by activation- (green) and information-based (red) analyses in this subject. Voxels marked by both techniques are shown in yellow. The sizes of these sets of voxels are related to each other in the left panel of **E** for this subject. The analysis has been performed with similar results for each of 11 subjects. The voxel set analysis for the group is shown in **E** (middle and right). **F** replicates the comparison map shown in D with one modification. Instead of using the FDR threshold obtained separately for each map, the FDR threshold obtained for the activation-based map (p=0.0013) has been used for both maps. As a result less voxels are marked in the information-based map and the expected FDR is lower than 5%. However, this does not entail a substantial qualitative change to the information-based map or the comparison. The slices shown in **A-D** and **F**, are slices 4, 6, 8, 10 and 12 (in anatomically ascending order from left to right) of 15 axial occipito-temporal slices acquired. The right side of each slice image represents the right hemisphere.

References

[1] Worsley, K. J., Evans, A. C., Marrett, S., Neelin P. (1992) A three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab.* **12**(6):900-18.

[2] Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C. & Evans, A. C. (1994) Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping* **1**, 214-220.

[3] Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D. & Frackowiak, R. S. J. (1995) Statistical parametric maps in functional imaging: a general linear approach *Human Brain Mapping* **2**,189-210.

[4] Poline, J. B., Worsley, K. J., Evans, A. C. & Friston, K. J. (1997) Combining spatial extent and peak intensity to test for activations in functional imaging. *NeuroImage* **5**, 83-96.

[5] Talairach, J. & Tournoux, P. (1988) *Co-Planar Stereotaxic Atlas of the Human Brain*. Thieme Medical Publishers, New York.

[6] Prüssmann, K. P. (2004) Parallel imaging at high field strength: synergies and joint potential. *Top Magn Reson Imaging* **15**(4): 237-44.

[7] Menon, R. S., Ogawa, S., Strupp, J. P. & Uğurbil, K. (1997) Ocular dominance in human V1 demonstrated by functional magnetic resonance imaging. *J. Neurophysiol.* **77**, 2780–2787.

[8] Menon, R. S. and Goodyear, B. G. (1999) Submillimeter functional localization in human striate cortex using BOLD contrast at 4 Tesla: implications for the vascular point-spread function. *Magn. Reson. Med.* **41**, 230–235.

[9] Cheng, K., Waggoner, R. A. & Tanaka, K. (2001) Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging. *Neuron* **32**, 359–374.

[10] Goodyear, B. G. and Menon, R. S. (2001) Brief visual stimulation allows mapping of ocular dominance in visual cortex using fMRI. *Human Brain Mapping* **14**, 210–217.

[11] Uğurbil, K., Toth, L. & Kim, D. S. (2003) How accurate is magnetic resonance imaging of brain function? Trends Neurosci **26**:108–14.

[12] Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L. & Pietrini, P. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* **293**, 2425-30.

[13] Edelman, S., Grill-Spector, K., Kushnir, T. & Malach, R. (1998) Towards direct visualization of the internal shape space by fMRI. *Psychobiology* **26**, 309-321.

[14] Cox, D. D. & Savoy, R. L. (2003) Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage* **19**, 261-70.

[15] Carlson, T. A., Schrater, P. & He, S. (2003). Patterns of activity in the categorical representations of objects. *J Cogn Neurosci* **15**, 704–717.

[16] Formisano E, Esposito F, Di Salle F, Goebel R. (2004) Cortex-based independent component analysis of fMRI time series. *Magn Reson Imaging* **22**(10):1493-504.

[17] Krzanowski, W. J. (1988) Principles of Multivariate Analysis: A User's Perspective. Clarendon Press: Oxford.

[18] Boynton, G. M., Engel, S. A., Glover, G. H. & Heeger, D. J. (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* **16**, 4207-21.

[19] Benjamini, Y. & Hochberg, Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* Ser. B **57**, 289–300.

[20] Genovese, C. R., Lazar, N. A. & Nichols, T. (2002) Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* **15**, 870–878.

[21] Kanwisher, N., McDermott, J. & Chun, M. M. (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* **17**, 4302-11.

[22] Epstein, R. & Kanwisher, N. (1998) A cortical representation of the local visual environment. *Nature* **392**(6676), 598-601.

[23] Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L. & Haxby, J. V. (1999) Distributed representation of objects in the human ventral visual pathway. *Proc Natl Acad Sci USA* **96**, 9379-84.

[24] Wager, T. D. & Nichols T. E. (2003) Optimization of experimental design in fMRI: a general framework using a genetic algorithm. *Neuroimage* **18**:293–309.

[25] Van Essen, D. C., Drury, H. A., Joshi, S. & Miller, M. I. (1998) Functional and structural mapping of human cerebral cortex: solutions are in the surfaces. *Proc Natl Acad Sci USA* **95**: 788-795.

[26] Dale, A. M., Fischl, B. & Sereno, M. I. (1999) Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage* **9**,179–194.

[27] Kriegeskorte, N. & Goebel, R. (2001) An efficient algorithm for topologically correct segmentation of the cortical sheet in anatomical MR volumes. *NeuroImage* **14**, 329-346.

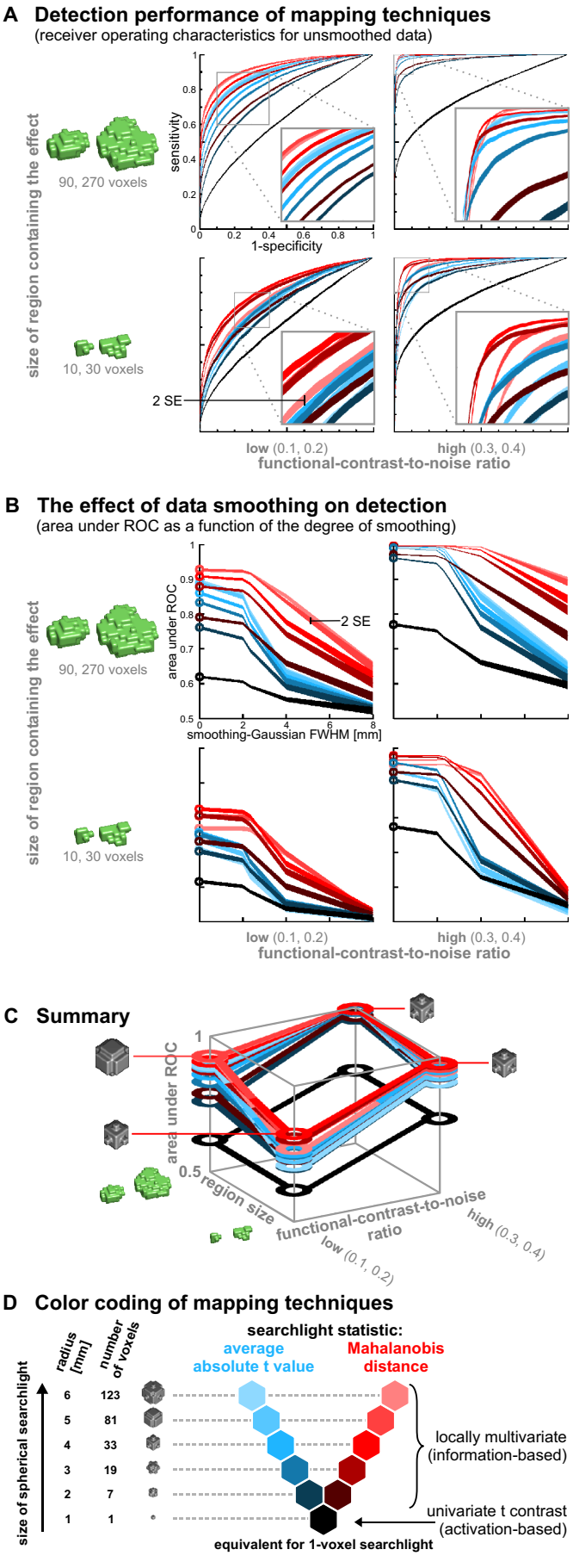

Figure 1

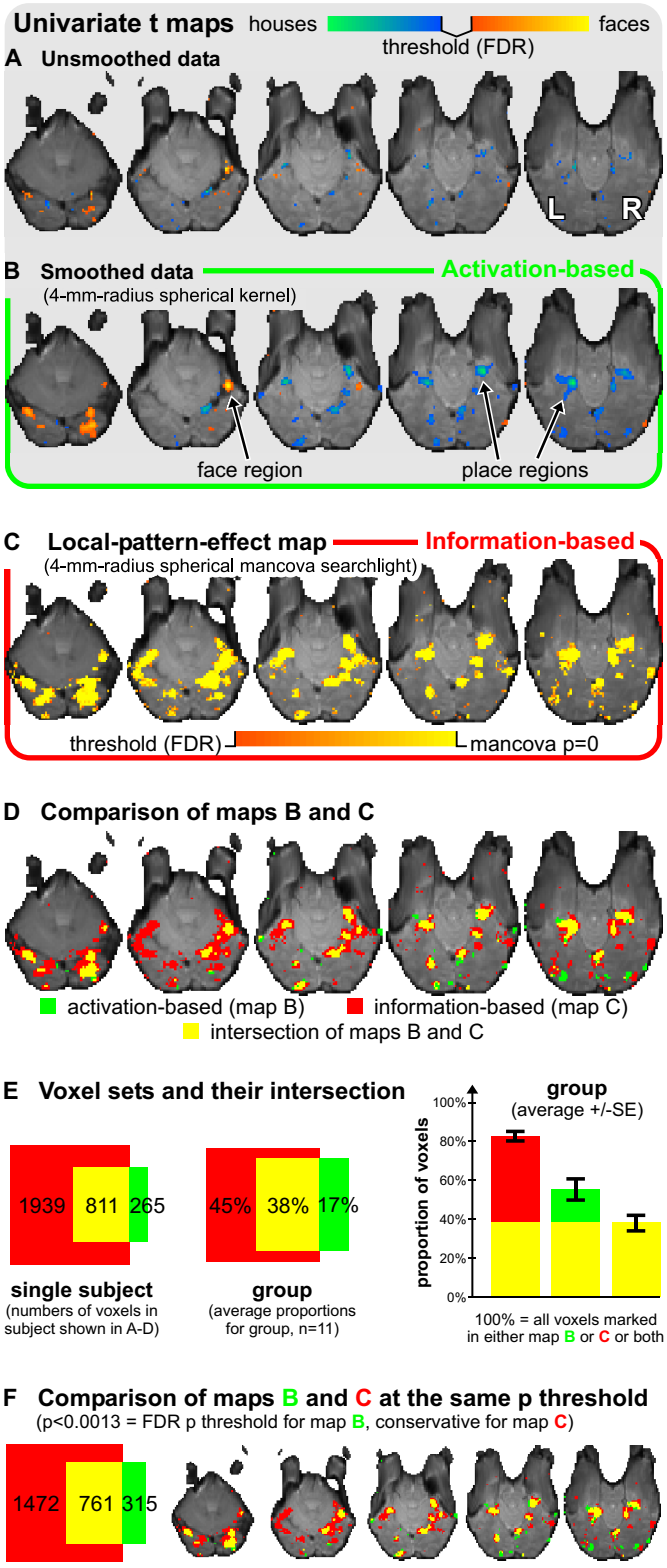

Figure 2